

## **Presentation**

**Sylviane CARDEY**

Coordinator, IM in NLP & HLT,  
Centre Tesnière, Université de Franche-Comté, France

This special number of the BULAG contains research project papers of special merit drawn from those written by students of the fourth intake (2011) of the European Erasmus Mundus Masters Course, IM in NLP & HLT - International Masters in Natural Language Processing and Human Language Technology (EMMC 2008-0083).

The papers have been written as a requirement of the International Masters students' research projects which they presented at the 4th International Workshop on Natural Language Processing and Human Language Technology, Universidade do Algarve, Faro, June 20-21, 2013.

The International Masters, which is a research masters, is administered by a consortium of four European universities: Université de Franche-Comté (coordinator), France, Universidade do Algarve, Portugal, Universitat Autònoma de Barcelona, Spain, and University of Wolverhampton, United Kingdom. The students study in 2 or 3 countries - one learns a language better if one practises it in the country where it is spoken. Furthermore, by means of its language policy which extends to the course of study, the course provides competence in more than one language, an essential for the Computational Linguist.

Being a research masters, the course content has input from the consortium's research findings. The theories developed and used by the different research centres in the consortium are varied – thus there is no attachment to some specific 'school' for this Masters and this avoids a closed view and ensures that the students are open minded. The papers collected here indeed reflect this 'open' policy. Each of the students, in conducting research as an integral part of the course programme, has been co-supervised by two teachers attached to research centres associated with two of the consortium members where they have studied.

The nine papers presented here have in common natural language processing and human language technology both from the theoretical point of view and also for varied applications. In respect of particular languages, aspects of variously Chinese, English, French, German, Italian, Khmer, Korean, Romanian and Spanish have been studied individually and in certain cases comparatively.

Elizabeth Rodríguez Sánchez's paper concerns the mathematical formalisation of microsystemic linguistic analysis.

Concerning the broad domain of text applications, Sovann En's paper concerns optical character recognition for Khmer, Luis Espinosa Anka's paper is concerned with definition extraction and Jacopo Ottaviani's with opinion mining.

Five papers have in common the domain of multilinguism and in particular machine translation; Daria Brazhnikova's paper is concerned with textual entailment across languages, Raluca-Sinziana Bujoreanu's with noun phrases in technical corpora, José Camacho Collados' with syntactic simplification, Oscar Mendoza's with collocation retrieval and finally Katherin Perez Roja's with film subtitles.